ナワバリをめぐるn人ゲーム(数理計画モデルにおける最適化理論)

AUTHOR(S):

寺岡, 義伸; 仲川, 勇二

数理解析研究所講究録
798巻 1992年 97-107

ナワバリをめぐる n人ゲーム

姫路工大理 寺岡義伸 (Yoshinobu Teraoka)
岡山理大工 仲川勇二 (Yuji Nakagawa)

1. 序論

ここで扱う問題は，生物進化の説明づけへの理論的道具としてゲーム理論が注目されることとなったモデル[2]，2種の動物が1つの縄張りをめぐって競うゲームから抽出された問題である。

出発点となった基本モデルは，次のようにまとめられる：

2人の players (Player I, II) は，価値Vを持った縄張りをめぐって対立している。この対戦に際して，各 player は誇示 (display)，挑み (escalate)，逃げ (retreat) の3つの行動がとれるものとし，この行動にもとづいて，次の2つの純戦略の中から一方を選択するものとする。

タカ戦略H：勝ち負けがはっきりするまで戦いを挑む。

ハト戦略D：まず誇示をし，相手が戦いを挑むようであれば，直ちに逃げ出す。

もし両者共戦いを挑むときは，いずれか一方が傷つき逃げ出すこととなる。そして敗れた方は，価値Vが入手できないだけでなく，価値Cをも失うものとする。そして，上記の戦略HとDに対して，以下のようにpayoffが定められる。

(i) I, II 共に戦略Hを選んだ時は，確率pでIが勝ち(IIが負け)，確率qでIが負ける(IIが勝つ)ものとする。ここに$p>0$, $q>0$, $p+q=1$である。

(ii) 一方のplayerが戦略Hを選び，他方が戦略Dを選んだ時は，戦略Hを選んだ方が価値Vを手に入れ，Dを選んだ方は価値Cを失うものとする。

(iii) I, II 共に戦略Dを選んだ時は，IとIIは$p:q$の比で価値Vを分け合うことにする。

2人のplayersはHかDのどちらの戦略を選ぶのが最適であるかを考えなければならない。このような仮定のもとで，以下のような利得双行列が得られる：

		II: H	II: D
I	H	$pV-qC,\ qV-pC$	$V \quad 0$
	D	$0,\ V$	$pV,\ qV$

上記の利得双行列に対する平衡戦略は容易に求まる[3]。この平衡戦略の1つに進化的に安定な戦略(evolutionanly stable strategy）略してESSと呼ばれるものがある。ESSの厳密な定義と数学的な構造についてはDammeの書物[1]第9章にくわしくふれられている。

2．縄張りの持久戦 — n人ゲーム

前節のモデルには時間的経緯が入っていなかった。しかしながら現実の場面にあっては，対戦は誇示ではじまるが，しばらくはにらみ合いが続き，このにらみ合いでどこまで誇示を持続できるかという問題が生じてくる。この場合，より長い時間持続できた方が勝ちとなる。ところが，この際誇示を持続するためには何らかのコストがかかるはずである。たとえば，生物進化学では繁殖をはじめるのが遅れるというコストが，ある新製品の市場獲得の競争では広告費用というコストが，考えられる。したがって誇示のコストは時間につき増加と仮定できる。このような状況のもとで，各playerにとって最適な持続時間を決めるのが，本報告の目的である。ここでは，payoffの形が全playersに共通となっているような非協力n人ゲームで取扱うことにする。

この問題を次のようにモデル化する：

n人のplayers（Player $1, \cdots, n$）は各々$[0, \infty)$内のどの時点まで頑張るかを決めようとしている。大きな時点を選ぐ

だ方が勝ちとなり，勝者は価値Vを敗者は価値0を得る．しかしながら 時刻 $t\in[0,\infty)$ まで持続するためには，各々 $h(t)$ のコストを費やさなければならない．ここにコストは

$$h(0)=0,\quad h'(t)>0 \quad \text{for } t\in[0,\infty),\ \text{さらに}$$

$$h(t)\nearrow\infty \quad \text{as } t\nearrow\infty$$

であると仮定する．次に n人のplayersのうち k人が最後まで頑張り同時刻で断念したときは，価値Vをk人で平等に分割するものとする。各playerは他の$n-1$人のplayersの持続時間を考えに入れた上での自分の最適持続時間を決めなければならない。

このn人ゲームにおいては，各playerにとっての競争相手は，他の$n-1$人というよりはむしろ，他の$n-1$人の中で最後まで頑張るplayerであろう。そしてplayerに与えられた情報様式によって2つの型の定式化が考えられる．全員が互に他の$n-1$人のplayersがまだ頑張っているのか もう既に断念してしまったのかが常に観測できる場合を，このゲームはNoisy型であるといい，逆に他の$n-1$人の行動が観測できない状態に置かれているとし，自分の予定した計画時間が実現されてみてはじめて，既に何人が断念してしまったのか あるいはまだ何人が頑張っているのかが知らされる場合を このゲームはSilent型であるという[3]．

3. Silent型の持久戦

n人のplayersの各々は 互に他の$n-1$人の行動を観測できない状態に置かれており，$(0, \infty)$のどの時点までねばるかを決定し，自分の決めたその計画時間が実現されてみてはじめて，既に何人が断念してしまっていたのか，あるいはまだ何人が頑張っているのかが知らされる ということだから，Player iの純戦略は$x_i \in [0, \infty)$とする。$i = 1, \cdots, n$.

$M_1(x_1, \cdots, x_n)$をPlayer iが純戦略x_iを用いたときのPlayer 1への期待利得とし，Player 2からPlayer nの選んだ純戦略$x_2, \cdots, x_n$の中で最大な値をyとする。そして，$x_1 = y$のときはPlayer 2からPlayer nの中の$k-1$人が最大値yを選ぶことになったとする。ただし$k = 1, \cdots, n$. そうすると

$$
(3.1) \quad M_1(x_1, \cdots, x_n) = \begin{cases} -h(x_1), & x < y \\ \frac{1}{k}V - h(x_1), & x_1 = y \\ V - h(x_1), & x_1 > y \end{cases}
$$

が得られる。混合戦略はゲームの対称性より全playersに共通なcdf $F(x_i)$とし，区間$(0, u) \subset (0, \infty)$上のdensity part $f(x_i) > 0$のみで構成される，というクラスから平衡戦略をみつけ出すことにする。点0でのmass partを0としてよいことは，2人ゲームの解析から容易にわかる[4].

Player 1が純戦略xを用い，他の$n-1$人が混合戦略$F(x_i)$を用

りたときの Player 1 への期待利得 $M_1(x, F, \cdots, F)$ は, Player 2 から Player n が選んだ純戦略のうち最大な値 y が

$$(n-1)f(y)\{F(y)\}^{n-2} \quad \text{for} \quad y \in (0, u)$$

で与えられる pdf をもつから, 次のようになる：

$$(3.2) \quad M_1(x, F, \cdots, F) = \begin{cases} V\int_0^x (n-1)f(y)\{F(y)\}^{n-2}dy \\ \quad -h(x)\int_0^u (n-1)f(y)\{F(y)\}^{n-2}dy, & 0 \leqq y \leqq u \\ \{V - h(x)\}\int_0^u (n-1)f(y)\{F(y)\}^{n-2}dy, & y \geqq u. \end{cases}$$

したがって, $M_1(x, F, \cdots, F) = v_1^*$ for all $x \in (0, u)$, および, $\int_0^u (n-1)f(t)\{F(t)\}^{n-2}dt = 1$ を解くと

$$(3.3) \quad (n-1)f(t)\{F(t)\}^{n-1} = h'(t)/V, \quad 0 < t < u$$

が得られ. (3.3) を (3.2) に代入すると

$$M_1(x, F, \cdots, F) = \begin{cases} [\{V - h(u)\}/V]h(x), & 0 \leqq x < u \\ [\{V - h(x)\}/V]h(u), & x \geqq u \end{cases}$$

が成立する. 以上より次の定理を得る.

定理 1. 次のような cdf $F^*(x)$ を考える：

$$F^*(x) = \begin{cases} \left\{\dfrac{h(x)}{V}\right\}^{\frac{1}{n-1}}, & 0 \leqq x \leqq u^* \\ 1, & x > u^* \end{cases}$$

ここに $u^* = h^{-1}(V)$ である.

そうすると, $(F^*, \cdots, F^*)$ は (3.1) の形で与えられる n 人非0和ゲームの1つの平衡点であり, この平衡点に対応す

る Player i への平衡値を v_i^* とすると

$$v_1^* = \cdots = v_n^* = 0$$

となる。▨

4. Noisy 型の持久戦

ここでは，前節とは逆に，各 player は互に他の $n-1$ 人のうちの何人がまだ頑張っているのか，あるいはもう既に断念してしまったのかが常に観測できるものとする。このモデルに対して，とりあえず次のような定式化を行うこととする。

Player i の純戦略を $x_i \in (0, \infty)$ とし，この純戦略に対して Player 1 への期待利得が次式で与えられるような n 人非0和ゲーム $M_i(x_1, \cdots, x_n)$ を考える：

$$(4.1) \quad M_1(x_1, \cdots, x_n) = \begin{cases} -h(x_1), & x_1 < y \\ \dfrac{1}{k}V - h(x_1), & x_1 = y \\ V - h(y), & x_1 > y. \end{cases}$$

ここに，y は Player 2 から n の $n-1$ 人が選んだ純戦略のうち最大な値であり，$x_1 = y$ となるのは Player 2 から n のうち $k-1$ 人が最大値 y を選ぶことになったとする。$k = 1, \cdots, n$.
このゲームも対称ゲームであるから，(4.1) で与えられる利得関数は全 players に共通であり，したがって混合戦略もまた全 players に共通と考えてよいから，$[0, \infty)$ 上の cdf $F(x_i)$ とし，平衡戦略としては，点 $x_i = 0$ での mass part は 0 で，$(0, \infty)$

上の density part $f(x_i)>0$ のみで構成される $F(x_i)$ の中からみつけることとする[3]. したがって, Player 1 が純戦略 x を用い, 他の $n-1$ 人が混合戦略 $F(x_i)$ を用いたときの Player 1 への期待利得を $M_1(x, F, \cdots, F)$ と書くことにすると, Player 2 から n が選んだ純戦略のうち最大な値 y は前節と同様に $(n-1)f(y)\{F(y)\}^{n-2}$ for $y \in (0, \infty)$ で与えられる pdf をもつとしてよいから

$$(4.2)\quad M_1(x, F, \cdots, F) = \int_0^x \{V - h(y)\}(n-1)f(y)\{F(y)\}^{n-2}dy + \int_x^\infty \{-h(x)\}(n-1)f(y)\{F(y)\}^{n-2}dy, \quad x \geqq 0$$

となる. そこで, $M_1(x, F, \cdots, F) = v_1^*$ for all $x > 0$ および $\int_0^\infty (n-1)f(t)\{F(t)\}^{n-2}dt = 1$ を解くと

$$(4.3)\quad (n-1)f(t)\{F(t)\}^{n-2} = \frac{h'(t)}{V}e^{-\frac{h(t)}{V}}, \quad \text{for } t > 0$$

が得られる. この pdf をもつ $F(\cdot)$ に対して

$$M_1(x, F, \cdots, F) = 0 \quad \text{for all } x \geqq 0.$$

そして 上記の cdf を F^* とすると

$$\{F^*(x)\}^{n-1} = 1 - e^{-\frac{h(x)}{V}} \geqq 0 \quad \text{for all } x \geqq 0$$

が得られて 次の定理を得る.

定理 2. 次のような cdf $F^*(\cdot)$ を考える:

$$(4.4)\quad F^*(x) = \left\{1 - e^{-\frac{h(x)}{V}}\right\}^{\frac{1}{n-1}}, \quad \text{for } x \geqq 0.$$

そうすると $(F^*, \cdots, F^*)$ は (4.1) の形で与えられる n 人非 0 和ゲームの1つの平衡点であり，対応する Player i への平衡値を v_i^* とすると $v_1^* = \cdots = v_n^* = 0$ となる。▥

ところで (4.1) の形の payoff では，他の $n-1$ 人中何人が既に断念したかということへの学習が入っておらず，最後まで $n-1$ 人を相手にすることを前提とした，しかし最後の player の行動時刻のみ観測できるとした戦略を導くものであった。そして，$F^*(\cdot)$ は，時々刻々のプレーヤの数の変化を組み込んだ cdf とはなっていないが，他の $n-1$ 人が全員途中経過が何であっても最初の参加者の数のみに依存する cdf $F^*(\cdot)$ を混合戦略として各 player が採用し，最後まで頑張る1人のみのみを問題にする，といった状態で，互に自己の期待利得を最大化するという意味での1つの平衡戦略となっている。しかも平衡値は 0 で維持されている。

そこで，$n-1$ 人中 m 人が残っているという状態で Player i の純戦略を $x_i \in (0, \infty)$ とし，この状態と純戦略に対しての Player I への期待利得を $M_1^m(x_1, \cdots, x_m)$，m players の選んだ純戦略のうち最大な値を y^m とすると (4.1) は

$$(4.5) \quad M_1^m(x_1, \cdots, x_m) = \begin{cases} -h(x_1), & x_1 < y^m \\ \dfrac{1}{k}V - h(x_1), & x_1 = y^m \\ V - h(y^m), & x_1 > y^m \end{cases}$$

のように書き換えることができる．ここに，m players がまだ残っているとき Player 1 は断念していないとする．

(4.1) と (4.5) を比較し，(4.4) を観察することにより，m players が残っている場合に対して，次のような cdf $F_{n-1}^{m}(\cdot)$ を support $[l_m, l_{m-1}]$ 上に定義する：

$$(4.6)\quad F_{n-1}^{m}(x) = K_m\left\{e^{-\frac{A(l_m)}{V}} - e^{-\frac{A(x)}{V}}\right\}^{\frac{1}{m}},\quad l_m < x < l_{m-1},$$

ここに l_m は，残りの競争相手が m 人になった時刻を示す．したがって

$$0 \equiv l_{n-1} \leqq l_{n-2} \leqq \cdots \leqq l_1 \leqq l_0 \equiv \infty$$

であり，さらに K_m は時刻 l_m までに残されている確率：

$$K_{n-1} = 1,\quad K_{n-2} = 1 - \left\{1 - e^{-\frac{A(l_{n-2})}{V}}\right\}^{\frac{1}{n-2}},\ \cdots$$

として定まるものとする．そうすると すべての m に対して

$$M_1^m(x, F^m, \cdots, F^m) = 0 \quad \text{for all } x \in (l_m, l_{m-1})$$

が成立する．

上記の考察を正確に忠実に定式化するには DP による再帰的関係の最適方程式を用いるのが効率的であるように考えられるが，ここでは (4.1) が残こり $n-1$ 人の最後に着目して定式化されたものであり，(4.4) が最後の 1 人に対して Noisy 型の平衡戦略となっていること，および 平衡値が 0 となる

一連の性質を利用した。

以上をまとめて 次の定理を得る:

定理3. l_mを残っている競争相手がmになった時刻とし, このmに対して cdf の部分 (4.6) で与えられる $F^m_{n+1}(x)$ を考えよ。そうすると この $F^m_{n+1}(x)$ で構成される cdf は (4.5) の形の payoff で構成される n人非0和ゲームの1つの平衡点となり, その平衡値は0である。

参考文献

[1] E. Damme: Stability and Perfection of Nash Equilbria, Springer-Verlag (1987).

[2] J. M. Smith: Evolution and the Theory of Games, Cambridge University Press (1982).

[3] 寺岡義伸: ナワバリをめぐる2人ゲーム, 京都大学数理解析研究所講究録 680「計画数学とその関連分野」275-284 (1989).

[4] 寺岡義伸: ナワバリをめぐる2人ゲームII, 京都大学数理解析研究所講究録 741「確率ゲーム理論とその周辺」76-91 (1991).